DOCUMENT RESUME

ED 087 546 PS 007 029

AUTHOR	Riley, Christine A.; Trabasso, Tom
TITLE	Comparatives, Logical Structures and Encoding in a Transitive Inference Task.
INSTITUTION	Princeton Univ., N.J. Dept. of Psychology.
SPONS AGENCY	National Inst. of Mental Health (DHEW), Bethesda, Md.
PUB DATE	73
NOTE	36p.; A paper based on this research was presented at the biennial meeting of the Society for Research in Child Development (Philadelphia, Pa., March 29-April 1, 1973)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Behavioral Science Research; *Cognitive Processes; *Concept Formation; *Elementary School Students; Feedback; Learning Processes; *Logic; Logical Thinking; *Preschool Children; Training
IDENTIFIERS	Relational Learning; *Transitive Inference Tasks

ABSTRACT

Four to five-year-old children were asked questions on length involving single or double comparatives in a transitive inference task. The number of comparatives varied within or across pairs. The pairs of sticks were color-coded. Initial pair-wise discrimination training on four adjacent pairs from a five stick array with only verbal feedback was followed by tests, without feedback, on all possible pairs. When both comparatives were used, subjects learned adjacent pairs faster and more often reached criterion than when only one comparative term was used in training. In testing, subjects were successful on transitive inference tests only when the double-comparative relation was used within pairs during training. Retraining with visual feedback increased the number of subjects solving the inference test only in this condition. The findings are discussed in terms of how many children may use processes other than logical operations to make transitive inferences. (Author/DP)

Comparatives, Logical Structures and Encoding in a Transitive Inference Task[1]

Christine A. Riley and Tom Trabasso[2]

Princeton University

Comparatives and Transitive Inference

Send proofs to: Tom Trabasso
Psychology Department
Princeton University
Princeton, New Jersey 08540

Abstract

Four-to five-year-old children were asked questions on length involving single or double comparatives in a transitive inference task. The number of comparatives varied within or across pairs. The pairs of sticks were color-coded. Initial pair-wise discrimination training on four adjacent pairs from a five stick array with only verbal feedback was followed by tests, without feedback, on all possible pairs. When both comparatives were used, Ss learned adjacent pairs faster and more often reached criterion than when only one comparative term was used in training. In testing, Ss were successful on transitive inference tests only when the double-comparative relation was used within pairs during training. Retraining with visual feedback increased the number of Ss solving the inference test only in this condition. The findings are discussed in terms of how children may use processes other than logical operations to make transitive inferences.

The present study examines some factors in a recent experiment by Bryant and Trabasso (1971) which led to successful transitive performance by four-to six year-old children. The importance of their finding bears re-examination since preoperational children, in theory (Piaget, 1960), are supposed to be unable to logically add the relationships that A is greater than B (A>B) and that B>C, resulting in the inference that A>C, using B as a middle term. Further, Bryant and Trabasso (1971) used procedures which appear to satisfy <u>all</u> of Smedslund's (1969) criteria for diagnosing a mental ability while at the same time, avoiding problems of using perceptual illusions, verbal reports or justifications which confound interpretation of transitivity results (Brainerd, 1973).

Bryant and Trabasso (1971) trained children (aged four through six years) to remember four length comparisons: AB, BC, CD and DE. The child first learned to discriminate each pair with feedback separately and in order. The discriminations were then relearned with the order of the pairs randomized. Following training, the children were tested without feedback on all ten possible pairs in a random order, allowing simultaneous measurement of the

(1) retention of critical adjacent pairs (pairs BC and CD)

(2) a transitive inference pair (BD) and

(3) end anchor effects (pairs involving A or E).

In response to the Braine-Smedslund controversy (cf. Smedslund, 1969). Bryant and Trabasso controlled for the problem of visually perceiving the difference in the BD comparison _at the time of test_ by making all the information and logical operations symbolic, i.e. the lengths were associated with colors and the child could use only the colors to make his choice in response to a question of which was longer or shorter throughout training and testing.

To control for what Smedslund (1969) called "non-transitive" hypotheses and to insure memory for the adjacent comparisons, Bryant and Trabasso used a procedure independently developed by Youniss and Murray (1970). There were four adjacent comparisons instead of two: A>B, B>C, C>D and D>E. The terms of the critical transitive comparison, B and D, as well as the middle term C, were thus equally often longer and shorter in adjacent comparisons prior to test.

In two experiments, one with visual feedback where the sticks were exposed after a choice in training and the other with verbal feedback where only the correct relationship was stated after a choice in training, were conducted. The proportion of correct responses on the transitive tests were above chance, ranging from 78 to 92 percent, and the degree of success on the BD tests was highly correlated with performance on the crucial adjacent pairs, BC and CD. (Lutkus & Trabasso (in press) have replicated the visual feedback results on retarded adolescents with MA's = 5 and 6, finding the correlations between BC and CD with BD performance of the order $\underline{r}$ = .70.) Their conclusion was that retrieval of the premises (ordered information on the adjacent pairs) was crucial to making transitive inferences and that a failure in memory for this information may underlie failure on transitivity tasks rather than logic or an inability to coordinate middle terms, the traditional interpretation.

There are three aspects of the Bryant and Trabasso (1971) procedures which we wish to examine here. In particular, their training procedure departed from traditional methods in studying transitivity by making explicit the <u>reversible</u> relation of the stick lengths in each pair, i.e. they required the children to answer <u>both</u> comparative questions. This procedure may have helped to bring about a fuller understanding of the comparative relation, overcoming the tendency to reduce comparatives to class labels (e.g., longer becomes long). That is, the child may have adopted an encoding of the relations A>B and B<A such that an ordered set A, B was more easily established. This would follow if the child already knows that length is transitive since once he orders A, B and B, C he need only to note the equivalence of B to order the set A, B, C and derive from this internal representation the inference that A>C. The equivalence of B may have been facilitated by identifying it with a unique color. Our experiments focus on the procedure of asking both comparative questions within the same pair.

The construction of an internal linear, possibly spatial array (cf. Trabasso & Riley, 1973; Trabasso, Riley & Wilson, in press; Potts, 1972; Scholz & Potts, 1973) which promotes transitive inferences may have been aided by the way Bryant and Trabasso presented and tested the pairs. Although size and position cues were controlled, the sticks were arranged in a display box so that location relative to the ends and distance between members of the pairs were correlated with a spatial array. That is, the end-anchored members A or E were always next to the end of the display box. The physical distance varied from a minimum of one inch for adjacent pairs to a maximum of nine inches between A and E. If the child uses spatial representations as apparently do adults (Huttenlocher, 1968), these cues could aid construction via end-anchoring and location in the array. In our present series of experiments, all such cues are removed by presenting the sticks next to each other for all pairs in both training and testing. Lutkus and Trabasso (in press) used this control in testing and found that performance was unaffected relative to Bryant and Trabasso's (1971) Ss.

Finally, visual versus verbal feedback appears crucial to the above discussion. Although Bryant and Trabasso (1971) found comparable results for both types of feedback, it may be that the spatial aids were critical for the verbal condition since the visual feedback promotes end-anchoring and distance through absolute lengths and size differences whereas verbal feedback does not lend itself to ease of locating end-anchors. That is, with visual feedback the child can identify the shortest and longest sticks in absolute terms and use size to assess distance between sticks since they are correlated. We therefore decided to train first with verbal feedback followed by retraining with visual feedback.

We focused on the comparative relational questions in three experiments where the logical structure of the information was the same but the questions and their distribution varied. In Experiment I, we followed the Bryant and Trabasso (1971) procedure of using both forms of the comparative ("Which is longer, A or B?"; "Which is shorter, A or B?" etc.) within each pair.

In Experiment II, we followed the traditional procedure of asking only one comparative question throughout training, e.g. "Which is longer, A or B? (B or C?) (C or D? (D or E?)" This experiment tests whether a failure to make a transitive inference results from the way in which pairs are encoded rather than a failure to coordinate via middle terms, a failure in logic. If children reduce ordered relations to classifications as Piaget (1928) notes, they should fail in training not testing. Such labelling produces: if A is longer than B, then A is long and B is not long and if B is longer than C, then B is long and

C is not long, which places B into mutually exclusive classes and causes a contradiction during training.

In Experiment III, both comparative questions are asked but they are asked across rather than within pairs (e.g.,"Which is longer A or B?", "Which is shorter B or C?", "Which is longer, C or D?" and "Which is shorter, D or E?"). This experiment tests whether or not a child needs to be forced to encode the reversible relationships using both comparatives within each pair or whether he needs only to be cued to use both forms. That is, would a child infer B<A when he has only learned A>B from the fact that he learned that C<B?

Experiment I

Method

Apparatus and Procedure. Twenty-five sticks (3-5-7-9-11-in. long in each of five colors (red, blue, yellow, green and white)) were used. A wooden box with five holes (1-3-5-7-9-in. deep) which allowed appropriate sticks to protrude 2-in. from the top was used in training. A similar wooden box with two holes (1-in. deep)was used during testing; test sticks were all 3-in. long. In training and testing, the S saw only two sticks at a time, in adjacent holes, and the presentation box was screened off so that no spatial or location cues from the box was available. This was done by use of a screen with an opening just wide enough to present two adjacent sticks and high enough to cover the box.

The Ss were run individually in three or four separate, approximately one half-hour sessions on successive days. In session one, each S was pretested for knowledge of color names and comprehension of comparative length terms (longer and shorter). The S was asked to name the color of each of the 3-in. sticks and then was shown four pairs of blue sticks of different lengths. The S was asked to select the longer stick in two pairs and the shorter stick from two other pairs. The S was then shown the training box, and was shown how sticks of different lengths (but the same color here) appeared the same when placed appropriately in the box.

All Ss succeeded in naming colors and answering comparative questions, and could explain why the sticks appeared to be the same length in the training box by pointing to "how far down inside the box" each stick went.

In session two, training with verbal feedback and testing without verbal feedback was carried out. Five sticks were used for each S during training. Each stick was different in length (A= 11 in., B=9 in., C=7 in., D=5 in. and E=3 in.) and color (white, yellow, red, green or blue). The length-color combinations were counterbalanced over Ss using a randomized Latin square design.

In training S learned the comparative relationships of four adjacent pairs of sticks: AB, BC, CD and DE. First, the four comparisons were trained separately and in order starting with the longest pair for half of the Ss with order of start counter-balanced over sex. The position of the sticks was random but occurred equally often over trials.

Upon presentation of a stick pair, S was asked, "Which stick is longer?" or "Which stick is shorter?". Questions were random but equal in occurrence over trials and over left-right positions. The Ss response was to state the correct color. All feedback during session two was verbal and of the form, "Right, the red stick is longer (shorter) than the blue stick," or "No, the red stick is longer (shorter) than the blue stick," for correct or wrong responses, respectively. The Ss were trained to a criterion of eight out of ten correct choices on each pair.

Then, a concurrent training procedure was used where all four pairs were presented in blocks of four trials, with pairs randomized within each block. The Ss were trained to a criterion of six consecutive correct choices on each pair or six correct trial blocks. If the S did not reach criterion in 100 trials (25 blocks) training was discontinued for that day. Training was resumed the next day beginning with separate training. If the S did not reach criterion in 200 trials he was not tested.

Testing followed immediately after criterion was reached in concurrent training. Testing was the same as training except that <u>no feedback</u> was given. Each S was tested four times on each of the ten possible color pairs. These ten pairs included the four direct comparisons used in training and six new, indirect comparisons (AC, AD, AE, BD, BE and CE). The four test questions on each pair included two longer questions and two shorter questions. For each form of the question the sticks were in both left-right positions. The 40 resulting test questions were randomized with the constraint that the same pair could not be presented on two consecutive questions. Half of the Ss started with the first question and half of the Ss started with the last question with the two groups counterbalanced on sex.

After the test questions were completed the box and screen were removed from the table and the E showed the S the five 3-in. sticks. The E said, "These sticks are the same colors as the sticks we talked about before. Do you remember which color was the color of the very longest stick?". When the S chose a stick, The E put it on the table. The E then said, "Which one of these colors was shorter than the stick on the table, but longer than the other sticks in my hand?". This questioning was continued until the S "seriated" the five sticks.

All Ss who completed training with verbal feedback were retrained with visual feedback in a final session. All training and testing procedures were identical to those used in session two except for feedback presentation. The E removed the sticks from the box and showed them to the S after a response. At the same time, the E gave the same verbal feedback as in session two.

Subjects. The Ss were 28 children, 11 boys and 17 girls, ranging in age from 4-1 to 5-0 (mdn = 4-8). (The sample for the three experiments included 16 Ss run in their school and 73 recruited by newspaper; of the latter 46 attended nursery school so that 70 percent of the total sample attended nursery school . The sample was largely upper-middle class and white). Eight Ss in Experiment I were replaced to ensure that 20 Ss reached training criterion prior to testing. Three girls failed to reach training criterion, and five others (one boy, four girls) refused to complete training. The median age of the 20 Ss who reached criterion was 4-7.

Results

The data from the 20 Ss who reached criterion in training serve as the basis for analysis in this section. The eight Ss who did not complete training are discussed when we compare results across experiments below.

Insert Tables 1 and 2 about here

Training.

Errors (and trial of last error)in training were each subjected to the analysis of variance where the within-S factors were training pairs (AB, BC, CD and DE), phase (separate versus concurrent) and type of feedback (verbal versus visual). Error rates were too low to allow for assessment of position and question effects independent of the sequence in which they occurred. Since errors and trial of last error were highly correlated and yielded identical significant effects, we report only the results on errors. The mean number of errors (and trial of last error) per training pair when the pairs were learned separately are in Table 1, and when learned concurrently, in Table 2. Experiment I results are shown in the first column of each table. In the analysis of variance, all main effects and interactions were significant by F tests ($p<.05$) except the pair by type of feedback interaction.

As can be seen in the tables, Ss, in Exp. I learned the separate pairs rapidly but were slower in relearning them together in the second, concurrent discrimination phase. This effect was stronger in the first training session with verbal feedback, producing the significant phase by type of feedback interaction. The pairs differed among themselves in difficulty. Individual comparisonsindicated ($p<.05$) that the pairs ordered themselves AB=DE>CD>BC with the largest effects occurring in the second, concurrent phase, especially when feedback was verbal, yielding the significant pair by phase and pair by phase by type of feedback interactions. Retraining with visual feedback reduced errors considerably.

The general pattern of these results resembles quite closely that found by Bryant and Trabasso(1971), particularly the serial position effects in the concurrent training, the general ease of learning in the separate training and the overall ease of training four concurrent discriminations.

Testing.

The primary data of interest in testing are the proportions of correct responses on test items. For convenience, these are presented in Table 3 in three groups: (1) the crucial inference test, BD, (2) the middle term pairs, BC and CD and (3) end-anchor pairs which involve A or E. The first two rows of Table 3 show the test results for Exp. I following training with verbal feedback and retraining with visual feedback, respectively.

Insert Table 3 about here

First we note that all the proportions for Exp. I are above the chance level of one-half since the 99 percent confidence bounds on this hypothesis with n=20 is 0.65. Although the observed proportion of 0.68 for the BD test is above chance, demonstrating that transitive inferences were made by Ss trained only with verbal feedback, that value is significantly lower than the 0.82 reported by Bryant and Trabasso (1971) in their verbal condition ($p<.05$). Furthermore, 12 out of 25 = 48 percent of their subjects were correct on all four transitive tests while seven out of 20 = 35 percent were perfect in the verbal condition of Exp. I.

The pattern of findings in testing are similar to those reported by Bryant and Trabasso (1971). Retrieval of the BC and

CD pairs was less than the level of performance on end-anchor pairs ($p<.05$) and was related to the degree of success on the BD pairs ($\underline{r} = .55$ $p<.05$). The performance following retraining with visual feedback was substantially increased for all pairs, especially on the crucial BC, CD and BD pairs. The latter showed significant increase of 20 percent ($p<.05$). There was no effect of question (longer vs. shorter) on the test questions ($p = .14$).

With the exception of the lower verbal performance, these results replicate the findings of Bryant and Trabasso (1971). The Ss were able to learn the pairs in training and give above chance correct answers to all test pairs, particularly the transitive test, BD, following both kinds of feedback. The reduction in our verbal condition on the BD test correlates with the average 28 percent loss in retention on the BC and CD pairs compared to a 12 percent loss for Bryant and Trabasso's Ss. That is, a difference of 14 percent in transitive responding is offset by a difference of 16 percent loss in retention of the information from the premises, BC and CD. The difficulty, then seems to be in storing and retrieving of the crucial information, underlining their contention that memory failure not logic, may account for failures in transitivity tests. Aside from differences in culture and the fact that all the British children were in school (but were of working class families), we suspect that Bryant and Trabasso's (1971) inadvertent use of spatial cues was critical in aiding the children to construct ordered, spatial arrays for use in storing the ordered set information and answering transitive questions. We have confirmed the spatial nature of these arrays elsewhere (Trabasso & Riley, 1973; Trabasso, Riley & Wilson, in press) but we have not studied the effect of environmental spatial cues which were present in the initial study by Bryant and Trabasso (1971).

Seriation. "Seriation" tests were given after each testing session. In the first, nine of the 20 Ss correctly seriated the lengths of the sticks in response to successive questions on the longest stick. The remaining 11 Ss misplaced three or more sticks, averaging 4.1 errors in placement. After retraining, the number who were perfect rose to 13; the remaining seven averaged 3.0 errors. These data are consistent with the other test results.

Experiment II

Method

Procedure. All materials and procedures in Experiment II were the same as those in Experiment I except for the kind of questions asked. In order to examine the hypothesis that Ss succeed in the task because both forms of the comparative question are used in training, only one form of the comparative was used in Experiment II to train the adjacent pairs. For the "longer" group, all the training questions were "Which stick is longer?"; for the shorter group, "Which stick is shorter?" In addition, all Ss were retrained with visual feedback in the third session, regardless of whether they reached criterion in the second session.

Subjects. The Ss were 10 boys and 11 girls, ranging in age from 4-0 to 4-11 (mdn = 4-5). One other boy failed to name the colors on pretests and was not included; one girl in the "shorter" condition refused to continue during training and was replaced. The remaining 20 Ss, (mdn age = 4-6) were randomly assigned to one of two conditions. There were 5 boys and 5 girls in each condition.

Results

Training with verbal feedback

All Ss learned the separate pair training phase rapidly and at about the same rate as those Ss in Experiment I under both kinds of feedback. However, in the second, concurrent, training phase, only seven of the twenty Ss reached criterion before 200 trials of training were given. The median age of those who succeeded was 4-5; for those who failed, it was 4-6, ruling out age as a factor here.

Insert Table 4 about here

Table 4 summarizes the comparisons between Experiments I and II on the number of Ss who reached criterion, failed to reach criterion or refused to continue. In Table 4 it can be seen that more than twice as many Ss succeeded in reaching criterion when both comparatives were used as opposed to one comparative

(z = 2.31, $p<.05$). One other difference in Table 4 is the fact that more Ss refused to continue in Experiment I than Experiment II, although this difference does not reach significance ($p<.10$).

These differences are critical in the assessment of the role of comparative questions on making transitive inferences. In Experiment II, many Ss during the concurrent training seemed totally baffled and told E that she was "crazy". Piaget's (1928) claim that children reduce a comparative to a class label (nominalization) is most clearly demonstrated by one child who reported: "You have two sizes of sticks, long and short, and you keep changing which ones are which." This suggests that Ss code the relation as "A is long, B is not long." If they also code "B is long, C is not long", a contradiction results and Ss cannot reach criterion on the initial pairs. In contrast, Ss in Experiment I refused to complete the procedures more frequently, expressing that they understood what they were supposed to do but stated that they had trouble remembering the pairs. These data suggest strongly that a major source of difficulty in traditional studies of transitivity is the use of a single comparative term. The failure of a child here would appear to depend more on how he codes the relations, not a failure in logic or coordination, once they are coded. The data also suggest that the use of both comparatives was successful in bringing about ordinal relations between members of a pair whereas using only one comparative promoted nominal relations.

The conditional mean number of errors (and trial of last error) for Ss who did reach criterion in the concurrent phase of training with verbal feedback is given for separate pair training in Table 1 and for concurrent training in Table 2. It should be noted that these Ss made significantly more errors (and took longer) to master the concurrent phase than those Ss who reached criterion in Exp. I (t (25) = 2.04, $p<.05$), a result consistent with our previous analysis and discussion. The error data was subjected to an analysis of variance where the within-S factors were pair and training phase. (Because different numbers of Ss were successful in reaching criterion under verbal and visual feedback conditions it was necessary to analyze the data for the two training sessions separately.) Both the pair and training phase main effects and the pair by training phase interaction were significant ($p<.05$). Individual comparisons ($p<.05$) indicated that the pairs ordered themselves DE<AB = CD<BC. It should also be noted that six of the seven Ss who did reach criterion were trained using "longer" as the comparative (z =2.34, $p<.05$), a result consistent with the idea that unmarked forms of comparatives are acquired earlier in development (Donaldson & Wales, 1970), pointing to still another, linguistic source of difficulty in transitivity studies, independent of logical abilities.

Retraining with visual feedback

In Experiment II, all 20 Ss were retrained with visual feedback. Six of the original seven who met criterion with verbal feedback, also made criterion here. Five additional Ss also succeeded, and as can be seen in Table 2, their error rates were higher than those for Ss in Experiment I, ($t(29) = 2.52$, $p<.01$). It should be noted, however, that some of these Ss reached criterion for the first time in this session. The number of Ss succeeding was independent of the comparative used: 6 vs. 5 for the "longer" and "shorter" questions, respectively. There were also no age differences for those who reached criterion versus those who did not (4-5 and 4-6, respectively).

An analysis of variance on error data showed the main effects of pair and training phase, and the pair by training phase interaction to again be significant ($p<.01$). Individual comparisons ($p<.05$) ordered the pairs AB = DE>BC = CD.

While retraining with visual feedback helped, the use of one comparative in training still led to fewer Ss reaching criterion and more errors for those who did than when both comparatives were used.

Testing.

The test results for the seven and eleven Ss who reached criterion in training with each kind of feedback in Experiment II are given in the third and fourth rows of Table 3. The upper 99 percent confidence bounds on the hypothesis that $p = .50$ are

0.72 (n = 7) and 0.68 (n = 11), respectively. With verbal feedback training, the proportions correct for three end-anchor pairs (AB, CE and DE) are above this bound. With visual feedback and retraining, all but one (BC) are above, and there was no significant difference in overall performance of those Ss tested for the first time and those Ss tested for the second time. Neither proportion correct on the transitive pair, BD, is significantly different from the 0.68 correct in Exp. I. The poorest performance was on retention of the BC and CD pairs and the overall performance is significantly worse than that for Exp. I with verbal feedback ($z = 4.07$, $p<.01$ for verbal training and $z = 1.75$, $p<.05$, one-tail for retraining with visual feedback). There were no significant differences between the two test conditions in Exp. II and there were no significant differences between questions asked within each condition.

Seriation.

After verbal feedback training and testing, no Ss were able to "seriate" the colors correctly; they misplaced an average of 3.9 sticks. After retraining and testing, five Ss were able to do so; the other six misplaced an average of 3.5 sticks.

Experiment III

Method

Procedure. The materials and procedures for Exp. III were the same as those in Exp. I except for the distribution of the questions over the pairs in training. In order to examine whether or not both comparatives must be present in order for Ss to encode the reversible relationship within a pair, Experiment III used both comparatives across but not within pairs. There were two groups and each group was trained on two pairs using only the question "Which stick is longer?" and two pairs only with "Which stick is shorter?" For one group, the pairs AB and CD were trained with the "longer" question and pairs BC and DE with "shorter"; the other group had the reverse question-pair assignment.

Subjects. The Ss were 20 boys and 20 girls ranging in age from 4-0 to 5-1 (mdn = 4-7). The Ss were randomly assigned to two groups of 10 boys and 10 girls each and were obtained from the same sources as Experiment I. All Ss completed all sessions.

Results

Training

The training results for Experiment III are summarized in the third columns of Tables 1 and 2. All 40 Ss learned both phases of training in both conditions with ease. An analysis of variance on errors in training, where the within-S factors are pair, training phase and type of feedback produced the same significant effects ($p<.05$) as in Experiment I with the exception that the main effect for pairs was non-significant. The interactions of pair by training phase and pair by training phase by type of feedback showed that a pair effect was present in the concurrent training with verbal feedback. An analysis of variance across experiments on errors in the concurrent training phase with verbal feedback was performed. Although the main effect of experiment was significant ($p<.05$), individual comparisons showed that Exp. III was not different from Exp. I.

Testing.

The test results for Exp. III are given in the fifth and sixth rows of Table 3. The upper 99 percent bound on the hypothesis of chance with $n = 40$ is 0.62. When training was with verbal feedback, all proportions on the tests are within the chance level. Following retraining, the end-anchor proportions are above chance but those on the crucial adjacent pairs BC and CD and the transitive test, BD, are not. The seriation results following testing after verbal feedback training were similar: only one S was able to seriate following with verbal feedback training; four retrained Ss were able to seriate by color. Clearly, transitive performance was lacking in this experiment.

While one may be tempted to conclude that preoperational children are unable to make implicit inferences on ordinal relations unless forced to do so through questioning, a simpler, alternative explanation is supported by the data. Although the logical structure of the task is the same throughout all experiments, the procedure in Exp. III was such that Ss could learn to answer questions on the training pairs with a simple response rule. For one group, A, C and E are always the correct answers; for the other B and D are always correct. If S learns these rules, *he does not need to learn anything about the length relationships*. During the test series, S can continue to apply these rules where one of the "right answer" colors is present; for other pairs, he simply guesses. If the S uses this strategy where the between questions and comparative is counterbalanced, he should be correct at the chance level, as observed. A reanalysis of the best data, breaking down the test questions into those on which S can apply his rule and those where he has to guess demonstrates that Ss were using this strategy. In the two feedback conditions, where the response rule selects the correct stick, the proportions were .79 and .92, respectively. When the rule selected the incorrect stick, these proportions were .35 and .43, respectively. These differences were highly significant by sign tests ($p<.01$). The performance on the two types of questions, when averaged, yields the chance levels shown in Table 3. Hence, the results of Exp. III demonstrate that Ss do not necessarily use the logical information if there is another, simpler way of solving the problem in answering E's questions.

General Discussion

These experiments indicate that the use of both comparative terms to query an ordinal relationship between a pair of items is critical in successful transitive performance by preoperational children. Performance is further enhanced if more training is given with visual referents. Memory for the ordered information in the premises would seem to be a critical factor in successful performance. The use of only one comparative term, as often occurs in studies of transitivity, is conjectured to lead young children to nominal rather than ordinal relations, and subsequent failure in learning to order the members of adjacent pairs. This source of difficulty in transitive tasks would seem to be independent of coordination of members via a middle term since it occurs in training and not testing. Finally, if S can find an alternative way of solving the task, he may learn nothing about relations, nominal or ordinal.

Youniss and Furth (in press) criticize the procedures used here since they omit certain controls that would show the importance of the middle term or "psuedo-transitivity". The question, as we see it, is not that certain controls were run etc. but whether the Bryant and Trabasso procedures measure "operational transitivity". From a Piagetian point of view, one must have such a measure if one is to verffy the model (e.g. by showing whether or not a preoperational child can perform such a mental operation.) That is, one has to rule out all other methods of problem solution since these would not constitute operational transitivity.

Our view is that such requirements beg the question in so far as theory is concerned. That is, one assumes the theory to be true and then devises tests that purport to show its presence or absence. Our view is that science is inductive and we prefer to infer the mechanism underlying the behavior from empirical evidence. This means that our interest is in any process that the child uses to solve a problem and that is why we have stressed the role of memory. It is doubtful that anyone could devise an experiment that would completely rule out language, perception, memory, etc. and reveal only logical competence.

We believe that, indeed, the children we have studied in these tasks do not use operational transitivity to solve the problem if one means by that term coordination of the members of the premises via a middle term at the time of testing. A process model which seems to capture some of the main features of how transitive inferences are made is now briefly mentioned and appears to be consistent with our data. The child would first anchor the ends of the scale, i.e., isolate sticks A and E. Then, under double comparative questioning, he would reversibly encode the terms, A>B and B<A, (D>E and E<D) allowing S to insert B (and D) into an internal array: A,B and D, E. The middle pairs are next ordered (B,C) and (C, D) and if S makes an identity match on colors, he can insert C into the array: A B C D E. Once this representation is complete, he can answer any relational question about the sticks via the color codes using the order information that have been built into the array.

The data of Exp. I are consistent with this model. The end pairs are learned fastest and all tests involving A or E are answered best. Visual feedback promotes end-anchoring since the longest and shortest sticks are displayed. In the verbal feedback condition, this isolation is more difficult since S must find those colors which have only one label associated with them and these are shared by other colors. The fact that BC and CD are both more difficult to learn and are lower in the proportion of correct responses on test is consistent with the notion that information from these pairs is used later in constructing the array.

Potts (1973) and Schulz and Potts (1973) in verbal recognition tasks on adults have independently offerred a model similar to ours based upon Huttenlocher's (1968) spatial imagery model for the three-term series problem (an adult version of the transitivity task). The essential point is that Ss coordinate information at the time of input into what appears to be a spatial array. Once S has integrated the information into an ordered, spatial array, he can more conveniently hold, store or retrieve this information rather than four (or more) separate ordered pairs, and answer the inferential questions by an internal scan of the array. A solution of this type does not require operational transitivity when S makes an inference by restating the pairs. His solution is not a formal statement of a transitive inference although some notion of transitivity is required in order to create an ordinal scale via integrating adjacent pairs into the serial array. Thus, if something like the above model is a correct description of a child's behavior in our task (Trabasso & Riley, 1973; Trabasso, Riley & Wilson, in press have evidence that it is), then the Bryant and Trabasso procedures are indeed irrelevant as far as Piaget's theory is concerned since they don't measure operational transitivity. We would hope, however, that this does not diminish their contribution to our knowledge of what children do do in reasoning about their environment.

References

Brainerd, C. J. Order of acquisition of transitivity, conservation and class inclusion of length and weight. Developmental Psychology, 1973, 8, 105-116.

Bryant, P. E. & Trabasso, T. Transitive inferences and memory in young children. Nature, 1971, 232, 456-458.

Donaldson, M. & Wales, R. J. On the acquisition of some relational terms. In Hayes, J. R. (Ed.) Cognition and the development of language. New York: John Wiley and Sons, Inc.,1970.

Flavell, J. H. The developmental psychology of Jean Piaget. New York: Van Nostrand Reinhold Company, 1963.

Huttenlocher, J. Constructing spatial images: A strategy in reasoning. Psychological Review, 1968, 75, 550-560.

Lutkus, A. D. & Trabasso, T. Transitive inferences in preoperational retardated adolescents. American Journal of Mental Deficiency, in press.

Piaget, J. Judgement and reasoning in the child. New York: Harcourt, Brace, 1928.

Piaget, J., Inhelder, B. & Szeminska, A. The child's conception of geometry. New York: Basic Books, Inc., 1960,

Potts, G. R. Information processing strategies used in the encoding of linear orderings. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 727-744.

Scholz, K. W. & Potts, G. R. Cognitive processing of linear orderings. Journal of Experimental Psychology, 1973, in press.

Smedslund, J. Psychological diagnostics. Psychological Bulletin, 1969, 71, 237-248.

Trabasso, T. & Riley, C. A. An information processing analysis of transitive inferences. Paper presented to the Eastern Psychological Association. May, 1973.

Trabasso, T., Riley, C. A. & Wilson, E. Spatial strategies in reasoning: a developmental study. In R. Falmagne (Ed.) Psychological Studies of Logic and its Development. New York: Erlbaum Assoc., in press.

Youniss, J. & Furth, H. G. Reasoning and Piaget: A comment on Bryant and Trabasso, Nature, in press.

Youniss, J. & Murray, J. P. Transitive inference with nontransitive solutions controlled. Developmental Psychology, 1970, 2, 169-175.

Footnotes

1. This work was supported by Research Grant No. 19223 from the National Institute of Mental Health, United States Public Health Service. We wish to thank Norma Johnson for her assistance in running subjects. A paper based on this research was presented at the Society for Research in Child Development Meetings, March, 1973.

2. Author's address: C. Riley, Department of Psychology, Princeton University, Princeton, N. J. 08540.

Table 1

Mean number of errors (trial of last error)
in training: separate pair training.

Verbal Feedback

Experiment	I	II	III
Pair	(n = 20)	(n = 7)	(n = 40)
AB	1.40 (3.05)	1.20 (2.70)	1.03 (1.68)
BC	1.10 (2.25)	1.00 (2.05)	0.80 (1.33)
CD	1.25 (2.25)	0.65 (0.75)	0.78 (1.03)
DE	1.10 (2.10)	0.65 (1.35)	0.75 (1.43)
Total	1.19 (2.41)	0.88 (1.71)	0.84 (1.37)

Retraining with Visual Feedback

Experiment	I	II	III
Pair	(n = 20)	(n = 11)	(n = 40)
AB	0.25 (0.60)	0.40 (0.90)	0.23 (0.63)
BC	0.30 (0.45)	0.30 (0.45)	0.28 (0.33)
CD	0.45 (1.20)	0.45 (0.75)	0.50 (0.55)
DE	0.15 (0.30)	0.10 (0.25)	0.23 (0.28)
Total	0.28 (0.60)	0.31 (0.59)	0.31 (0.76)

Table 2

Mean number of errors (trial of last error) in training: concurrent training.

Verbal Feedback

Experiment	I	II	III
Pair	(n = 20)	(n = 7)	(n = 40)
AB	2.55 (6.85)	7.28 (16.14)	2.60 (5.90)
BC	5.35 (10.85)	12.29 (24.43)	3.53 (8.85)
CD	4.10 (9.35)	8.57 (18.86)	3.83 (7.90)
DE	3.10 (7.30)	3.43 (9.57)	3.05 (6.55)
Total	3.78 (8.59)	7.89 (17.25)	3.25 (7.30)

Retraining with Visual Feedback

Experiment	I	II	III
Pair	(n = 20)	(n = 11)	(n = 40)
AB	0.15 (0.15)	0.36 (0.82)	0.43 (0.75)
BC	0.90 (1.45)	2.18 (4.73)	0.43 (0.70)
CD	0.90 (1.95)	2.83 (5.18)	0.60 (1.03)
DE	0.35 (0.55)	1.27 (2.18)	0.33 (0.50)
Total	0.58 (1.03)	1.65 (3.23)	0.45 (0.74)

Table 3

Proportion of correct responses on test pairs (retention tests are starred).

Experiment	Feedback	n	Transitive Inference Pair	Critical Pairs		End-anchored pairs				
			BD	BC*	CD*	AB*	AC	AD	AE	BE
I	Verbal	20	.68	.65	.79	.93	.81	.75	.85	.80
	Visual	20	.88	.81	.91	.99	.90	.91	.93	.88
II	Verbal	7	.66	.54	.57	.75	.71	.61	.57	.61
	Visual	11	.72	.59	.68	.82	.82	.77	.77	.75
III	Verbal	40	.60	.54	.53	.61	.50	.59	.59	.56
	Visual	40	.64	.55	.61	.69	.63	.72	.65	.60

Table 4

Proportion (frequency) of subjects reaching criterion in the concurrent phase of training.

	Reached criterion	Failed to reach criterion	Refused to continue
Experiment			
I	.714 (20)	.107 (3)	.179 (5)
II	.333 (7)	.619 (13)	.048 (1)